Software Engineering Institute

Analyzing Cases of Resilience Success and Failure—A Research Study

Julia H. Allen
Pamela Curtis
Nader Mehravari
Andrew Moore
Kevin Partridge
Robert Stoddard
Randy Trzeciak

December 2012

TECHNICAL NOTE
CMU/SEI-2012-TN-025

CERT® Program

http://www.sei.cmu.edu

Carnegie Mellon

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

This report was prepared for the

SEI Administrative Agent
AFLCMC/PZE
20 Schilling Circle, Bldg 1305, 3rd floor
Hanscom AFB, MA 01731-2125

DM-0000050

Table of Contents

List of Figures

List of Tables

Acknowledgments

The authors thank the following individuals for their contributions to this research study:

- four collaboration partners who participated by contributing examples of disruptive events for research team analysis
- John Haller from the CERT Infrastructure Resilience Team and Lisa Young from the CERT Cyber Resilience Center for their assistance in formulating the follow-on research project strategy described in Section 4
- William Scherlis, SEI chief technical officer (acting), for his review of and guidance on the proposal for the follow-on research project that resulted from this initial study
- reviewers of this report (and earlier draft materials)
  - William Claycomb, CERT Enterprise Threat and Vulnerability Management Team
  - Linda Parker Gates, SEI Acquisition Support Program
  - David Zubrow, SEI Software Engineering Process Program

The authors thank David White, technical director of the CERT Cyber Resilience Center, for his continuous support and encouragement. The authors also acknowledge Richard Caralli, director of the CERT Cyber Enterprise Workforce Management Directorate and Summer Fowler, program manager, for their leadership in establishing the grant research project program that served as the funding source for this research study, as well as for their continuous support and encouragement.

Abstract

Organizations that are using the CERT® Resilience Management Model and organizations that are considering using it want information about the business value of implementing resilience processes and practices, and how to determine which ones to implement. This report describes the SEI research study that begins to address this need. It includes a discussion of the completed phase 1 study and a proposed phase 2 project. Phase 1 included forming a hypothesis and set of research questions and using a variety of techniques to collect data and evaluate whether resilience practices have a discernible (measurable) effect on *operational resilience*—that is, an organization's ability to continue to carry out its mission (provide critical services) in the presence of operational stress and disruption. The outcomes of phase 1 provide the foundation for the proposed phase 2. The longer term goal includes developing a quantitative, validated business case for prioritizing and implementing specific resilience practices, including decision criteria for selecting and measuring investments in improved resilience.

1 Introduction

Since the release of the CERT® Resilience Management Model [Caralli 2010, Caralli 2011], organizations that are considering using the model and organizations that are implementing it have sought information about the business value of using resilience processes and practices and how to determine which ones to implement. This same return-on-investment concern has been raised by the community at large for years when considering any process improvement effort. We, as a model-developing community, have generally not been able to provide a satisfactory response to questions of this type. This is particularly the case for new models that have yet to be broadly adopted and, thus, for which there are very few experience reports and case studies. This research effort was conceived to begin to remedy this situation.

In January 2012, the CERT Program's[1] Cyber Enterprise Workforce Development Directorate decided to allocate internal funds to begin to address organizations' questions through an initial research study titled "Analyze Cases of Resilience Success and Failure." A proposed follow-on research project (see Section 4.0) will make progress in evaluating whether resilience practices have a discernible (measurable) effect on *operational resilience*—that is, an organization's ability to continue to carry out its mission (provide critical services) in the presence of operational stress and disruption. In the longer term, we also intend to develop a quantitative, validated business case for prioritizing and implementing specific resilience practices, including decision criteria for selecting and measuring investments in improved resilience.

Throughout this report, we use the term "phase 1" to describe the initial research study performed from January through September 2012. We use the term "phase 2" to describe the work we propose as a follow-on to the initial study, subject to available SEI research and customer funding.

During phase 1, the research team explored the hypothesis and research questions described in Section 1.2 and formulated the revised hypotheses and research questions, described in Section 2.1. The team has collected a small sample of case data (disruptive events) from collaboration partners. We have used this data and other insights to develop a research scope and analysis approach that serves as a foundation for phase 2 of the research project in fiscal year (FY) 2013 (October 2012 through September 2013).

The remainder of this report is organized as follows:

- Section 1.1 describes the scope of phase 1.
- Section 1.2 presents our initial research objectives, hypothesis, and research questions formulated during phase 1.
- Section 2 describes the general research method for exploratory and explanatory research, the phase 1 data collection and coding approach, and several initial analysis methods based on mixed-methods research (qualitative and quantitative).

[1] The CERT Program is part of Carnegie Mellon University's Software Engineering Institute (SEI).

- Section 3 presents our research observations and lessons learned from phase 1, including a general description of our collaboration partners, the cases they provided, and insights we gained from working with this case data.
- Section 4 provides an overview of our proposed FY13 phase 2 project.
- Section 5 summarizes the report.

Terms used throughout this report are defined in the glossary.

1.1 Scope

For phase 1, the initial scope of candidate resilience practices are those described in the CERT® Resilience Management Model (CERT-RMM), Version 1.1 [Caralli 2011]. CERT-RMM is a capability-focused maturity model for process improvement that reflects best practices from industry and government for managing operational resilience across the domains of security management, business continuity management, and aspects of information technology (IT) operations management.[2] CERT-RMM defines operational resilience as

> *the emergent property of an organization that can continue to carry out its mission in the presence of operational stress and disruption that does not exceed its limit.*[3]

Stress and disruption arise from the realization of operational risk: failed internal processes; failures of systems or technology; the deliberate or inadvertent actions of people; and external events (such as natural disasters) [Caralli 2011]. To expand, operational resilience is the organization's ability to protect and sustain high-value services and their associated assets (information, facilities, people, and technology such as systems and software) to achieve the service mission. An operationally resilient service is one that can meet its mission under times of disruption or stress *and* can return to normalcy when the disruption or stress is eliminated. A service is *not* resilient if it cannot return to normalcy after being disrupted, even if it can temporarily withstand adverse circumstances [Allen 2011a].

Practices in the model focus on improving the organization's management of key operational resilience processes. These improvements enable high-value services to meet their mission consistently and with high quality, particularly during times of stress and disruption [Caralli 2011].

The primary resilience requirements that must be met in the presence of and after disruption are confidentiality, integrity, availability, and privacy. The applicability of a specific type of resilience requirement varies depending on the asset type, as shown in Table 1.

---

[2] For more information about CERT-RMM, see the book titled *CERT® Resilience Management Model: A Maturity Model for Managing Operational Resilience* and the CERT Resilience Management Model pages on the CERT website [Caralli 2011, CERT 2012].

[3] *Disruption* in this definition applies to a disturbance that does not exceed the service's operational limit. A catastrophic loss of infrastructure would not be considered a disruption.

Table 1: Extension of Resilience Requirements to Types of Resilience Assets

Resilience Requirement	Asset Type			
	People	Information	Technology	Facilities
Confidentiality	—	x	—	—
Integrity	—	x	x	x
Availability	x	x	x	x
Privacy	—	x	—	—

The research team selected availability of technology assets as the scope for phase 1. Our rationale for this scope was driven by the wide range of operational disruptions (caused, for example, by denial of service, disruption of service, jamming, losing communication nodes) that can occur in an operational theater when technology is not available. In addition, our rationale was based on an initial sample of case data from collaboration partners where availability of technology assets was disrupted. The Department of Defense (DoD) Cyber Science and Technology Priority Steering Council Research Roadmap [King 2011] provided additional impetus by identifying time to restore operational capability and effort to restore operational capability as two key metrics of interest for characterizing a resilient infrastructure.[4] We used these metrics as our primary focus and sought to identify recommended practices that measurably reduced time and effort to restore operational capability in the presence of and after an attack.

1.2 Objectives, Research Questions, and Hypothesis

The section describes the objectives, research questions, and hypothesis for phase 1. Revised research questions and hypotheses for the phase 2 project are described in Section 4.1.

1.2.1 Objectives

The goal of this overall research effort is to inform the general question "*Do resilience practices have a discernible (measurable) effect on operational resilience?*" This objective can be expressed as a statement: *Measure the contribution of resilience practices to reducing the occurrence and impact of disruptive events*. One could argue that there is ample anecdotal evidence and some documented case studies in the resilience domains described above showing that improved practices do have a discernible and positive effect on an organization's ability to be more resilient. That said, the authors of this report are currently unaware of comprehensive research results that evaluate and demonstrate a measurable effect that can also be used as a basis for selecting and prioritizing which resilience practices to implement. We are exploring a wide range of sources to validate this observation and to place this research within the relevant research literature and communities of practice. Several of these sources are cited in Section 4.6.

[4] The Roadmap measures time to restore operational capability in minutes, hours, and days, and in effort to restore operational capability in number of cyber specialists required to resolve a significant attack: 100, 10, 1, or automated.

Resilience practices include any actions, methods, or techniques that help satisfy a resilience requirement. One example pertinent to our technology availability scope is a service level agreement that specifies availability requirements for key servers. Resilience practices that address this requirement could include redundant equipment, regular backups, and the ability to restore from backups in a specified time frame.

We study the general question by investigating and analyzing resilience cases—cases in which an organization was successful or unsuccessful in satisfying a resilience requirement in the presence of and after a disruptive event. While the handling of any specific disruptive event may have both successful and unsuccessful aspects, we intend to identify these distinct aspects in different resilience case descriptions. Thus, any resilience case can be identified to have one of the following:

- a positive outcome where the impact of the disruptive event is no more than some critical threshold (as defined by a resilience requirement)
- a negative outcome where the impact of the disruptive event is more than some critical threshold (as defined by the resilience requirement)

Figure 1 depicts concepts and relationships associated with operational resilience that are relevant for this research effort, including the relationship of resilience requirements to impact thresholds in the context of services, assets, and controls. Resilience processes and practices implement the protection and sustainment strategies for high-value services and assets. In addition, this figure describes how operational resilience is driven by enterprise requirements. It is derived from the comparable Figure 2.6 in the *CERT Resilience Management Model: A Maturity Model for Managing Operational Resilience* [Caralli 2011]; thus it serves as the overall context for this research effort (both phase 1 and phase 2).

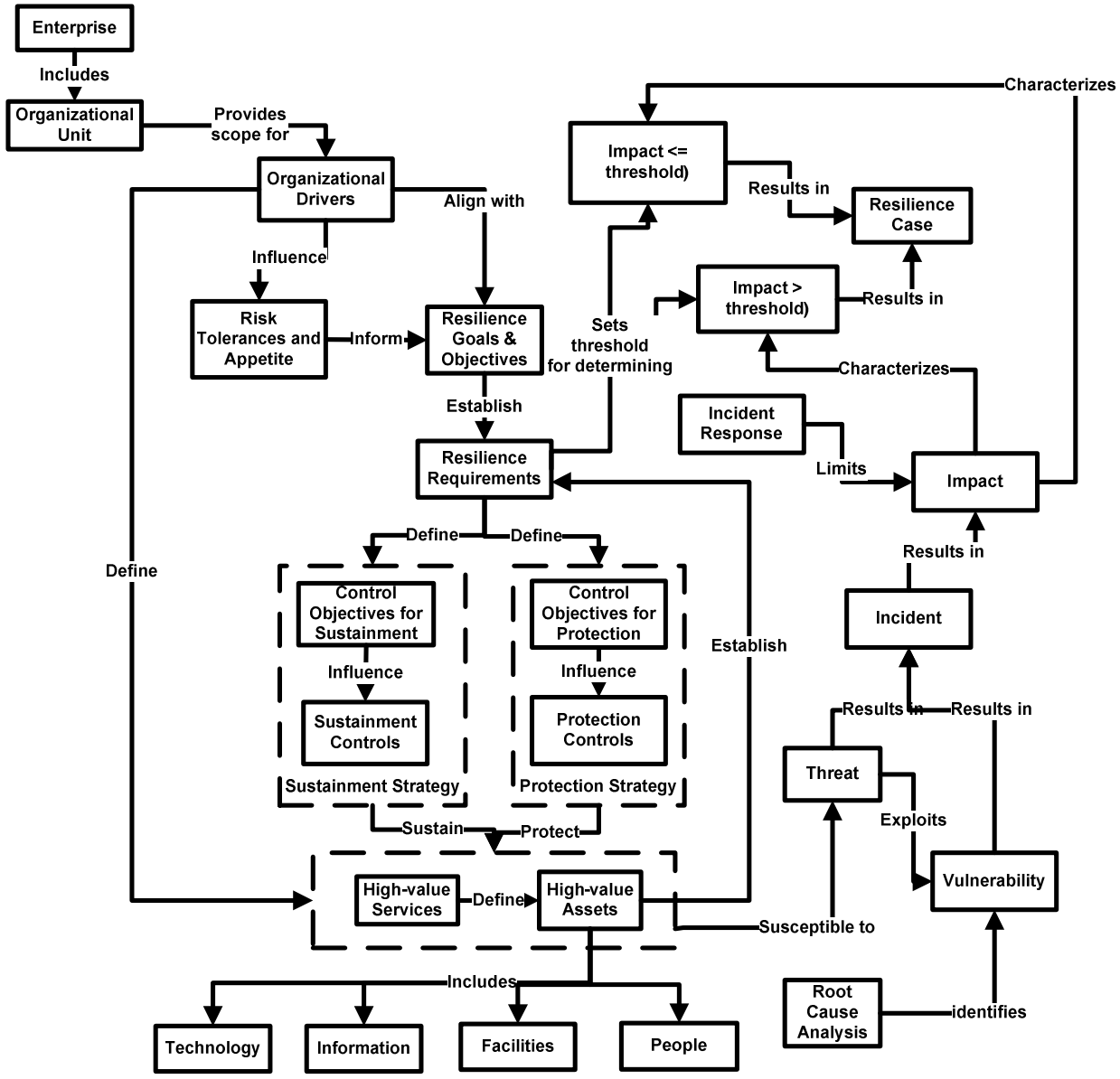

Figure 1: Relevant Concepts and Relationships Associated with Operational Resilience

We initially focused on two impact measures, the time to restore service and the effort to restore service, reflecting the interests of the Office of the Secretary of Defense, Science, and Technology (OSD S&T) [King 2011] and the emphasis on technology availability.

1.2.2 Research Questions

We further elaborated on the broader question "*Do resilience practices have a discernible effect on operational resilience?"* by addressing two more specific research questions in addition to those noted above:

1. for disruptive events (incidents) with a positive outcome (impact <= threshold): What was the contribution of resilience practices to the successful handling of the incident causing the impact? How much did the use of resilience practices contribute to success? (**Note:** A successful response also includes cases where the incident was thwarted and no impact occurred.)
2. for resilience cases with a negative outcome (impact > threshold): What was the source of the failure and the root cause of the impact? Would the organization have been successful if it had implemented (selected) resilience practices?

To identify gaps in and improvements to CERT-RMM resilience practices, we intend to explore the following additional research questions:

- Did our collaboration partners perform any resilience practices that they found to be useful and that are not included in CERT-RMM?
- Are there practices suggested by CERT-RMM that don't have much positive effect and, therefore, could be reduced or eliminated—specifically those that have a high cost?
- Are the time and effort to restore reduced as a result of implementing sets of related resilience practices?
- Did partners implement any particular practices only because they were suggested by CERT-RMM and were effective? We would eventually like to make a claim on this question, supported by evidence and analysis.

1.2.3 Hypothesis

Phase 1 focused on the following general hypothesis:

> *Disruptive events are more likely to have a positive outcome (i.e., response considered successful) when resilience practices are implemented than when they are not.*

The research objectives, questions, and hypotheses, and the types of data we needed to explore them resulted in the research approach described in the next section. Additional and more detailed hypotheses are also discussed.

2 Research Approach

2.1 Overview

This work was largely framed as an exploratory research study applying the multiple (or comparative) case study method described by Robert Yin [Yin 2009]. Thus, we did not start with concrete, specific hypotheses other than the general one stated in Section 1.2. One goal of the exploratory phase 1 study was, in fact, to generate more specific hypotheses for the phase 2 explanatory research project.

We defined the following analysis outcomes for phase 1 but accomplished only two of them, as indicated, primarily due to the lack of sufficient resilience case data:

- the generation of more specific hypotheses (accomplished)
- evidence of the validity of those hypotheses (deferred)
- a level of confidence associated with that evidence (deferred)
- a statement of the nature and degree of generalization of the research results (deferred)
- an understanding of experiments and studies that can be conducted and important data to be collected that will help establish or refute specific hypotheses in phase 2 (accomplished)

We did formulate several more specific hypotheses that we intend to explore during phase 2. The percentages in these hypotheses are examples; they might be modified according to the confidence goals of our research partners.

- Confidence that 80% of incidents in the category of interest (technology availability, information confidentiality) are reported.
- A known list of factors explain at least 80% of the time to handle a specific category of incidents
- A known list of factors explain at least 80% of the cost of handling a specific category of incidents.
- A known list of factors cause at least 80% of a specific category of incidents.
- A known list of factors may control the reduction of a specific category of incidents by at least 80%.
- Identified factors are analogous to or inform the selection of resilience practices that confirm these hypotheses.

In addition to the hypotheses that were focused on resilience cases and their underlying state of resilience practices, we formulated questions associated with the practical measurement of factors directly tied to resilience practices, cases, and impacts. These questions articulate the need to better understand the measurement scaling of such factors and the degree of their repeatability and reproducibility.

- What are the potential factors to be measured to study the research questions?
- What is the proper measurement scale for each factor?

- What is the proper analytical or modeling approach related to each factor or combination of factors?
- Can these factors be measured practically, reliably, and with acceptable degrees of repeatability and reproducibility?
- Many of these questions can be answered using small-scale, statistically designed experiments [Montgomery 2012] and studies, as described in Section 4.4, thereby enabling the collection of the correct types of data to be used in analyzing cases. Answers to these questions will help us perform a more targeted and quantitative analysis subsequent to phase 1.

While phase 1 was limited to exploratory research, in phase 2 we will strive to explain relationships between resilience practices and organizations' incidents through hypothesis testing [Sheskin 2011]. This overall research effort is, therefore, an example of mixed-methods research [Creswell 2011], which involves both qualitative and quantitative research. As Creswell and Clark describe, mixed-methods research may be appropriate when investigators do not know the exact questions to ask, variables to measure, and theories to guide the study, possibly because of the newness of the research topic. This is the case for phase 1. Creswell and Clark state that "in these situations, it is best to explore qualitatively to learn what questions, variables, theories, and so forth need to be studied and then follow up with a quantitative study to generalize and test what was learned from the exploration." Mixed methods are ideal for this type of research. The qualitative analysis conducted in phase 1 paves the way for quantitative analysis in phase 2.

2.2 Data Collection and Coding

The research team reached out to potential collaboration partners and collected case data using the definition of a resilience case as defined by the Resilience Case Request for Data template (see the appendix). For phase 1, we selected cases on a "target of opportunity" basis by engaging with collaboration partners where a strong trust relationship existed between individuals. Many of the individuals we have approached for case data thus far are reluctant to share such sensitive data and, in some cases, unwilling. However, where there is an existing customer project work statement, a CERT-RMM licensing agreement, a CERT-RMM Users Group relationship, or a professional relationship between peers characterized by mutual respect and trust, we find that individuals are more willing to share data and are able to provide sufficient rationale to their internal stakeholders to obtain permission to share the data. To date, using the form in the appendix, we have obtained a description of eight cases from four organizations. We were not able to obtain all the information for all fields for all cases but did obtain sufficient data to explore the analysis approaches described in this section.

The market sectors represented by the four collaboration partners that kindly provided case data for phase 1 included U.S. federal civilian agencies, U.S. defense contractors, commercial service providers, and academia. The eight cases we collected from these four partners described the following general types of incidents:

Business continuity

- power outage requiring migration to a backup site
- hurricane requiring migration to a backup site
- fire requiring migration to a backup site

Technology availability

- web server compromise due to malicious software
- email phishing resulting in user credentials being provided to a malicious website
- prevention of a zero-day exploitation that would have installed malicious software
- insufficient monitoring and alerting to determine if requests to download unauthorized files were correctly blocked
- filled log volume preventing user access to email

Coding of cases is a critical process whereby information gathered through case review and interviews is entered into a resilience case template and (ultimately) a resilience case database according to a prescribed methodology documented in a codebook. Due to insufficient data, we did not pursue a defined case coding approach (other than the use of the template by research study members) or the development of a resilience case database during phase 1. We did perform in-depth analysis of our research questions and the case data template to ensure that we were collecting sufficient data to address our research questions. In the process of performing this analysis, we identified several gaps,[5] which will be remedied in phase 2. All these tasks will be explored in more depth during the follow-on project.

After performing the preliminary analysis on a single event described in Section 2.3, we determined that the analyses necessary to support phase 1 and phase 2 require a range of events of a similar type over time (at least 7 with similar resilience practices; optimally 30–40). We have identified a source for such data where a deep trust relationship exists with the partner providing the data. We plan to pursue this data collection approach during the conduct of phase 2.

2.3 Analysis Methods

Given the small data set we had to work with, we explored the question "How do I structure data to inform small controlled experiments to achieve a reasonable understanding of how to analyze resilience cases to inform our research questions?" Of the eight cases we collected, only one provided sufficient data to explore candidate analysis methods. Using that case, we experimented with two types of analysis: attribute agreement analysis and a preliminary conceptual Bayesian Belief Network (BBN). Our analysis experiments are described below.

2.3.1 Attribute Agreement Analysis

Attribute agreement analysis can be used to collect and assess ratings provided by subject matter experts (e.g., a minimum of 4 subject matter experts) and then to examine such ratings (e.g., a minimum of 20 item ratings) for agreement and divergence. We defined and executed the following variation of an attribute agreement analysis process to conduct an experiment on our ability to render reproducible, subjective judgments to code resilience cases. This process was performed by three members of the research team.

[5] Of particular note is the examination of "success" or evidence of resilience. The definition of what should compose resilience evidence does not scale across all events, preventing direct comparisons. The ambiguity also affects other desirable measures, such as the contribution of resilience practices to success and the threshold of failure.

1. Develop a detailed chronological time line of each action taken (or actions not taken that should have been) in the case. Identify CERT-RMM processes and practices that are relevant for each action. Prepare a spreadsheet that maps each action to the applicable CERT-RMM processes, goals, practices, and subpractices.[6] For example, one action was to notify the organizational computer incident response team when the suspicious event was first detected. Two of the CERT-RMM practices that apply here are IMC:SG3.SP1, Define and maintain incident declaration criteria, and IMC:SG4.SP3, Communicate incidents. For this case, we identified 18 actions and 30 applicable CERT-RMM practices (with specific subpractices).
2. Answer and score the following two questions for each CERT-RMM practice:
   - Question 1 (Q1): To what degree was this practice implemented in this case? Use a scale of zero to 10 (zero being not implemented at all; 10 being fully implemented) and do one of the following:
     a. Provide a brief rationale for each answer to Question 1.
     b. Specify that additional data is needed to answer this question and the nature of that data.
   - Question 2 (Q2): Given the answer to Q1, what is your subjective assessment of the role that this practice as implemented played in the resilience outcomes for phase 1? (Resilience response <= threshold; resilience response > threshold, expressed as time to restore and effort to restore.) Use a scale of -10 to +10, where -10 is a significant negative role (made the response much worse), zero is neutral (no role), and +10 is a significant positive role (made the response much better). Also, do one of the following:
     a. Provide a brief rationale for each answer to Q2.
     b. Specify that additional data is needed to answer this question and the nature of that data.
3. Initially, to ease the case analysis workload, we discussed answering Q2 only for practices that have answers at the extremes for Q1 (i.e., 0–2 or 9–10), as we may only care about analyzing practices that are significantly absent/weak/ineffective or significantly strong.
4. We initially evaluated each practice independently even though some practices may relate to others. We chose to defer analysis of practice interrelationships.
5. Practices for which we could not answer Q1 and Q2 were eliminated from consideration for this first experiment.

An answer to Q1 or Q2 is called a judgment. We needed 100 judgments to conduct our first analytical experiment (attribute agreement analysis). So if we could identify 12–13 practices in a case for which we could answer both questions (24–26 judgments) and four people could do this analysis, we would have approximately 100 judgments.

We observed that this analysis process was best done by two researchers first, independently, followed by their reconciling these first two sets of results. Then, one or two additional

[6] We found that we needed to go to the subpractice level to gain sufficient understanding and detail to perform a reasonable mapping. In the future, we may also need to map to finer grained control descriptions as identified in the CERT-RMM Crosswalk [Partridge 2011].

researchers performed the same analysis using the first set of results as their baseline. We determined that it was too difficult to reconcile three to four sets of analysis results simultaneously.

After three team members attempted this analysis, we realized that we lacked sufficient data to form a consensus, which led to subjective and varying results. We also determined that this approach was too resource intensive (in time and effort) to scale to the number of cases we would need to analyze to produce meaningful results. As a next step, we decided to focus on a subset of CERT-RMM practices that we believed provided the strongest contribution to a successful outcome and used the analysis approach described in the next section.

2.3.2 Conceptual Bayesian Belief Network

The purpose of this analysis was to develop a preliminary, conceptual Bayesian Belief Network (BBN). A BBN is a probabilistic graphical model that represents a set of random variables and their conditional dependencies.[7] We believed a BBN structure could be useful in helping us identify key practices and their interdependencies toward predicting resilience outcomes of time and cost for a specific case. Historical use of BBNs in the field of quality and reliability enabled a simplification of a model, with direct causal practices connected to performance outcomes and additional "upstream" practices providing a leading indication of the directly causal practices.

A research team member selected 19 CERT-RMM practices (from the 30 identified in the case above) that we believed would most likely predict a successful outcome if performed adequately. These practices were placed into a spreadsheet structured to evaluate cause and effect relationships among practices. All practices were listed as rows and columns in the spreadsheet. We then scored the intersection of each practice with every other practice as follows:

Score	Meaning
0	no relationship between these two practices
1	possible relationship between these two practices (weak influence)
2	agree that there is a relationship between these two practices (moderate influence)
3	strongly agree that there is a relationship between these two practices (strong influence)

For example, there is a strong relationship (a score of 3) between IMC:SG3.SP1, Define and maintain incident declaration criteria, and IMC:SG2.SP4, Analyze and triage events (assign disposition). There is no relationship between IMC:SG3.SP1, Define and maintain incident declaration criteria, and IMC:SG4.SP3, Communicate incidents (identify relevant stakeholders).

After we established and reviewed these relationships, we added the scores by practice and identified those with the highest score. Practices with the highest row scores indicated those that were most likely to be a precursor or input to other practices, and practices with the highest column scores indicated those that were downstream and more dependent on other practices. For example, VAR:SG3.SP1, Manage exposure to vulnerabilities, had the highest precursor (row) score. MON:SG2.SP3, Collect and record information, had the highest dependency (column) score. From this analysis, we developed a preliminary, conceptual BBN, attempting to identify

[7] According to http://en.wikipedia.org/wiki/Bayesian_network

practices that, if performed adequately, would most likely predict a successful outcome. However, this attempt at using a BBN did not immediately resolve to a simple, hierarchal model in which some practices were more directly causal of resilience outcomes and others were more indirectly causal. Instead, a high degree of bidirectional correlation was noted among the practices, confirming the need for more case data and in-depth analysis. The spreadsheets and diagrams resulting from this analysis are available upon request.

2.3.3 Data Types, Analysis Approaches, and Outcomes

At this point in phase 1, we decided to step back from individual cases and describe the types of data that would inform our research questions and hypotheses. We then described each data type's value, limitations, and challenges; the types of analysis we could do with each type of data; and candidate analysis outcomes. The results of this effort are shown in Table 2.

Table 2: Data Types, Candidate Analysis Approaches, and Potential Analysis Outcomes

Type ID	Data Type	Value/Limitation/Challenge	Analysis Approach/Method	Analysis Outcomes
D1	Individual events/cases[8]	Limitation: Single data point; difficult to conclude anything about practice efficacy from one data point	Attribute agreement analysis, Gage Repeatability and Reproducibility (R&R) and/or Cronbach's Alpha for the repeatability and reproducibility analysis; simple Pareto analysis	Use single event or case to be separately analyzed and coded by a number of domain experts for analysis of repeatability and reproducibility; Pareto list of most impacting or significant sustainment practices for the event or case
D2	Range of events[9] of a specific type[10] over a specific time frame[11]	Value: Likely most valuable Limitation: Conclusions relevant only for a given event type; need at least 5–7 items in each sample to conduct the hypothesis test	Hypothesis testing to compare groups of similar events or cases, or to compare practice implementations or significance	Establish data sufficient to analyze for systemic patterns or trends at the practice level as well as for repeatability and reproducibility purposes
D3	Experiments using environments such as XNET; could also include broader simulation[12]	Value: Observation in application	Fractional, factorial, statistically designed experiments in which only a small fraction of the total combinations of factors must be evaluated	*Significance* of individual practices or grouping of practices using p-values from experiments; statistical regression equations *predicting* resilience outcomes based on knowledge of factors and practices; follow-on experimentation can provide *optimization* guidance with regards to practice implementation and other factors
D4	Long-term, deep-dive trust relationship with a specific organization (pilot partner)[13]	Value: CERT-RMM appraisal experiences/results, tracking incidents as they occur, recommending improvements, tracking results after improvements, observation in application (likely most valuable)	Hypothesis testing across time as noted above; experimental analysis as noted above; time series and other probabilistic modeling; Monte Carlo and discrete event simulation; Weibull analysis for univariate time-bound data such as time to diagnose, time to recover	Rich comparisons, predictions, statistically guided diagnosis, dashboards

[8] Such as the cases we currently have

[9] Information confidentiality case data may fall into this category.

[10] For example, compromises of customer-facing websites or the restoration of IT operations following a disruption of continuity

[11] 30 days, 90 days, 6 months, 1 year

[12] The leap here is to the actual operational environment and the relevance of simulation results to it.

[13] Two candidates identified

Type ID	Data Type	Value/Limitation/Challenge	Analysis Approach/Method	Analysis Outcomes
D5	Survey data on a specific topic	Challenge: Identifying questions and survey community that would be willing to answer them (self-assessment)	Hypothesis testing of changes in responses to individual questions across time; comparison of responses to different questions at any point in time; trend analysis of answers across time; predictive modeling of questions aligned with outcome notions based on the state of affairs of questions aligned with leading indicators	Standard survey data analysis with statistical analysis for answering difference or trend questions
D6	Challenge workshop	Challenge: Identifying key stakeholders and information we are seeking from them, with likelihood of obtaining it	Affinity maps of stakeholder inputs; conjoint analysis of stakeholder beliefs and preferences; Pareto analysis; Strengths, Weaknesses, Opportunities, Threats (SWOT) analysis	New ideas on measures, analytics, and reporting; lessons learned
D7	Data from existing repositories, third parties such as vendors	Challenge: Getting data that is relevant to CERT-RMM practices	Hypothesis testing, experimental analysis, structured equation modeling of multivariate data; probabilistic modeling; different forms of simulation (systems dynamics, discrete, Monte Carlo)	Relationships (cause and effect, correlation) between practices, factors, etc.

From this analysis, we determined that to successfully conduct this research, we needed a range of incidents (between 7 and 100) of a comparable type (such as violations of information confidentiality) over a period of time (weeks, months). We identified a partner with the potential to provide such data and started to develop a guiding scenario and supporting analysis approaches with this data source in mind. We suspended further analysis of the existing set of eight incidents and formulated the data analysis research approach for the phase 2 project described in Section 4.4. This concluded the analysis task for phase 1.

3 Observations

Our primary observations are described throughout this report and summarized here. We also include additional observations and insights from the conduct of phase 1.

3.1 Defining Phase 1 and Phase 2

It is important to allocate ample time early in the research effort to define key terms, formulate clear research objectives, and identify research questions and hypotheses that support these objectives. This may seem obvious, but it is not easy. These activities required numerous iterations over the course of phase 1. Here are examples of some of our early discussion questions:

- What constitutes a resilience case?
- How are success and failure defined?
- Which cases are worth collecting and which are not? What are meaningful case selection criteria?
- Which cases are practical and affordable for the research team or the collaborating organizations to collect?
- What kind of case data should be collected, and how much of that data over what period of time or frequency is needed?
- What is the acceptable mix of subjective versus objective case data?
- What is the required quality and repeatability of the data to enable analysis and modeling?
- How should the analysis be structured to produce repeatable results? And to stand up to review and scrutiny over time?
- What sources of SEI, CERT, and external expertise can we draw upon?
- What should we use as our guiding references and sources (such as the work of Yin [Yin 2009] and Creswell and Clark [Creswell 2011])?
- Are we conducting explanatory research, exploratory research, or both?
- Are we seeking qualitative or quantitative methods and results, or both types?

When formulating a research proposal, think about validation methods, measures of success, and impacts of the research first; that is, begin with the end in mind. Include anticipating and mitigating threats to both external and internal validity, and then develop companion guiding scenarios and solution approaches.

In cases where we are unable to obtain true measures of impact (as is the case for unauthorized disclosure of sensitive information), it may be viable to identify surrogates for true impact, such as a reduction in both the number of such incidents and the costs of handling them. Early drafts for review need to describe research methods in the context of the overall goals for the research effort so reviewers have adequate context for judging the adequacy and appropriateness of those methods. Also, researchers should elaborate the range of methods needed to address the variety of research questions of interest.

3.2 Collecting and Coding Cases

Early on, we relied heavily on the CERT Program's experience in collecting and coding insider threat cases and derived significant benefit as we developed the Request for Resilience Case Data template included as the appendix. However, we identified a substantive difference in this research effort's objectives. Insider threat case analysis, in large part, is focused on characterizing the threat and identifying practices and controls that may help mitigate it. The derivation of insider threat controls has largely been an informal process relying on expert knowledge of the effectiveness of controls against various security vulnerabilities. Phase 1 focused on rigorously evaluating the efficacy of resilience practices in the presence of a realized threat (i.e., a disruptive event). While this area of focus is also of great interest for the insider threat research, we were unable to obtain relevant experience or learning from that body of work to inform phase 1.

It is important to ensure that the data being collected will adequately inform the research questions and hypotheses. This, too, may be obvious; but this task does require in-depth examination and analysis.

We discussed several approaches for coding cases in a reliable and repeatable manner. We also discussed several approaches for building a repository of resilience cases that would support some form of structured and automated analysis, including the use of tools. Based on the case data we were able to collect, we determined that these discussions and any decisions resulting from them were premature.

Many of the individuals we approached for case data were reluctant to share such sensitive data and, in some cases, unwilling even with existing non-disclosure agreements. However, where there was an existing customer project work statement, a CERT-RMM licensing agreement, a CERT-RMM Users Group relationship, or a professional relationship between peers characterized by mutual respect and trust, we found that such individuals were often willing to share such data. That said, in several instances, they were not able to provide sufficient rationale to their internal stakeholders to obtain permission to share the data beyond a summary level. Though this may not be a major issue when dealing with government organizations with which we have a formal relationship, we need to provide more information and assistance to help our private sector partners develop this rationale.

3.3 Analyzing Data

In parallel with working with collaboration partners to obtain case data, we drafted an approach for analyzing the data and then sought the review of two senior SEI researchers. Based on their feedback, we learned that it was premature to develop such an approach in the absence of case data and in the absence of a full description of the background and motivation for phase 1.

CERT-RMM is a management model, which means that the resilience practices are stated at a fairly high level when viewed from an implementation and action perspective. We found that we needed to go to the subpractice level to gain sufficient understanding and detail to perform a reasonable mapping of case actions to resilience practices. In the future, we may need to map to finer grained control descriptions as identified in CERT-RMM Crosswalk [Partridge 2011].

After performing the preliminary analysis on a single event, we determined that the analyses necessary to support phase 1 and phase 2 require a range of events of a similar type over time (at least 7 with similar resilience practices; optimally 30–40). Trends over time provide a much richer data set than individual events.

Here are some of the challenging analysis questions we discussed:

- If we know the impact of a given incident, how do we determine the contribution (or not) of a given practice?
- Is there practice correlation or causation with respect to incident impact?
- Do we have cases we can compare where incidents were handled well versus not handled well?
- Which practices tended to be more significant from a statistical viewpoint in determining different outcomes?
- Should we focus on individual practices, or are practice interrelationships and groups of practices more applicable?
- How do we determine the degree of practice implementation (say, on a scale of 1–10)?
- How do we need to record data and what do we need to record to make our analysis outcomes repeatable and defensible?
- Given the often sparse and incomplete case descriptions we had to work with, how reliable and repeatable are scoring and analyses performed by subject matter experts?
- Should we assess at the case level or the individual practice level? For example, how well did the organization detect and respond across the entire event? What allowed them to succeed or fail (greater focus on the root cause)?
- Should we consider opportunities to work with collaborating organizations and establish continuous data collection over a moderate to longer period of time?
- Can assessment data inform this research?

If we use a particular set of data to formulate candidate hypotheses, we need new and comparable data to confirm or refute those hypotheses. One approach is to analyze existing data and create a model, followed by a time period of collecting new data to use for validation of the model. Another approach, assuming sufficient case data, is to collect data and use half of it to analyze and develop a model, and then use the other half to immediately validate the model.[14] In this latter approach, there is no need to wait for a set period of time to collect more data for model validation purposes.

[14] In general, if you have a relatively small data set for which new cases are not readily available, you will probably want to split the data set into an analysis set and a test set. The analysis set is used to determine the effectiveness of a given set of practices for a specific security concern. The test set is used to test whether the hypothesis generated using the analysis set holds for the test set, thus giving evidence as to its generalizability. There is no set rule on how you choose the sizes of the analysis and test sets. It depends on your relative concern with generalizability versus hypothesis generation, all other things being equal. So, for example, if you are able to collect data describing 20 events, you should probably analyze 10 and save 10 for testing.

3.4 General

Other research teams that depend on incident case data may find the analysis approaches and methods considered during phase 1 and planned for during phase 2 to be of value.

The CERT Program Cyber Enterprise Workforce Management Directorate provided initial funding to support phase 1. This gave our research team the ability to explore our research questions, experiment with a variety of data and analysis approaches, and make key discoveries along the way as to how to more effectively conduct case-based research tied to resilience practice outcomes. Phase 1 was essential for identifying potential collaboration partners, identifying sources of case data, and developing a defensible proposal for phase 2.

4 FY13 Phase 2 Research Project Overview

4.1 Objectives and Scope

For our proposed FY13 phase 2 project, we intend to continue to pursue the objective of measuring the contribution of resilience practices to reducing the occurrence and impact of disruptive events. Phase 2 is the next step toward a larger, long-term objective of developing a method to evaluate the extent to which resilience practices contribute to an organization's ability to continue to carry out its mission in the presence of operational stress and disruption.

Due to the nature of the incident data available to us, as described in Section 2.2, the research initially will focus on information confidentiality—specifically, the unauthorized disclosure of sensitive data. We will examine the effect of resilience practices for preventing and containing data disclosure on reducing the time and cost (impact measures) to handle disclosure incidents. A set of practices will be proposed to and vetted by our research partners.

4.2 Research Hypotheses and Questions

The same general hypothesis is posed for phase 2 as the one we began with in phase 1: *Disruptive events are more likely to have a positive outcome (i.e., response considered successful) when resilience practices are implemented than when they are not.* Data will be collected and experiments conducted to test the more specific hypotheses that were formulated during phase 1 (listed in Section 2.1).

In testing the general hypothesis, we will examine each practice for its relative contributions as follows:

- to reducing the occurrence and impact of disclosure incidents
- to reducing the time and cost of handling disclosure incidents
- to the root causes of incidents
- the ratio of the cost of implementing the practice to the resulting cost savings of handling disclosure incidents

In addition to the research questions listed in Section 1.2, during phase 2 we will seek answers to these:

- What factors are driving the occurrence of disclosure incidents, and what controllable factors exist to reduce the rate of disclosure incidents? Might resilience practices and measures [Allen 2010, Allen 2011a, Allen 2011b] predict candidate factors?
- What factors contribute to the time required to handle disclosure incidents, and what percentage of the total time do they account for?
- What factors contribute to the cost required to handle disclosure incidents, and what percentage of the total cost do they account for?

Cost calculations will be informed by the Cost of Incidents and Mean Cost of Incidents metrics from the CIS Consensus Security Metrics [CIS 2010].

4.3 Guiding Scenario

Our guiding scenario, the unauthorized disclosure of sensitive information, is shown in Figure 2, as a time line describing potential events and the resulting impact. A qualitative description of the impact is provided, based on the classification of the disclosed data. In reality we find that the activities within the incident that realize this impact are variable and possibly unknown. In our time line, we assume the realization of the impact.

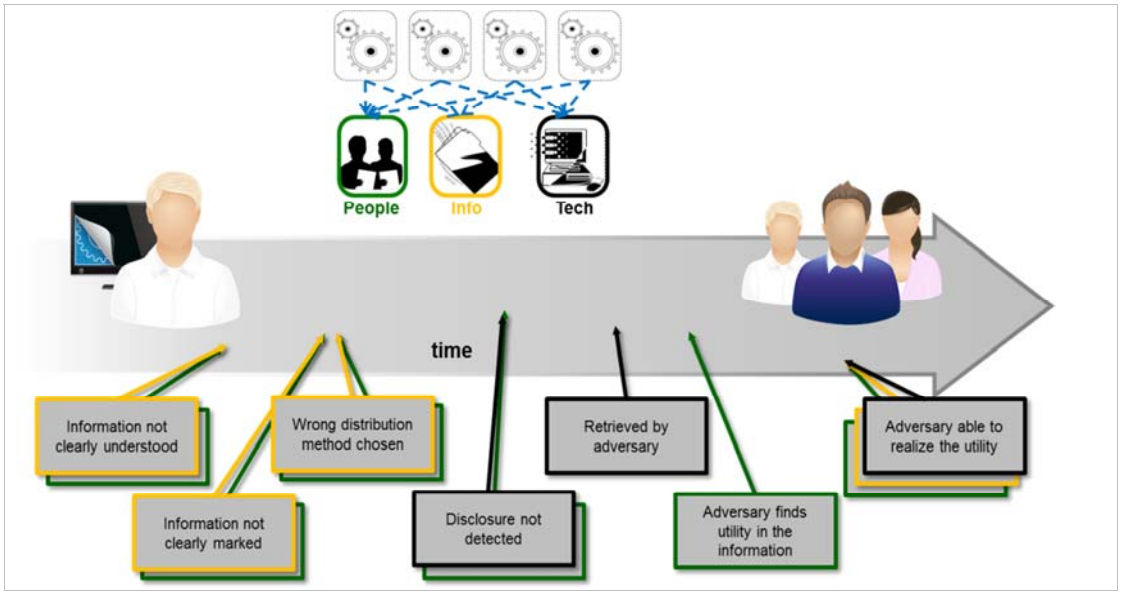

Figure 2: Unauthorized Disclosure of Sensitive Information

The precipitating activities involve two asset types: people and information (marked as green and yellow borders, respectively). People may not clearly understand the nature of the information or the impact of possible disclosure. The information may not be clearly categorized and may not be clearly marked. As the time line progresses, actions shift from protecting the information (or not) by people to protecting and using information through technology (marked as a black border).

If the information is well understood and clearly marked, the resilience practices for protecting information are considered adequately implemented. The time line progresses only if the people or technology practices are ineffective or missing. Unauthorized disclosure of sensitive information occurs when the user chooses a data distribution method inappropriate for the information's classification. Either the actions of people or an implemented technology practice could detect certain methods of unauthorized disclosure. Often this disclosure is reported by the originating user or the recipient of the data. Practices known as Cross Domain Solutions (CDS) are used to monitor for data marked for the wrong domain. These practices have various limitations. If the data bypasses these practices and their companion controls, the only remaining barrier to realizing the impact that results from disclosing sensitive information is the capability of the adversary and environmental factors.

The adversary's ability to retrieve sensitive information is dependent upon the environment in which it resides—its technological container such as a server, a website, or an accessible email attachment. The adversary must have access to this container. Once in possession of the information, the adversary must be capable of recognizing its utility. Finally, the adversary must

have the ability to use the information. This portion of the time line is dependent on the adversary's practices associated with effectively using people, information, and technology assets. Mitigation is dependent on the disruption of the time line at any of the described points, including those that affect the adversary's ability to use the information.

4.4 Data Analysis Approach

Phase 2 will focus on the measured contribution of resilience practices to the reduction of the occurrence and impact of disruptive events. Our research team will use a quasi-experimental research design commonly referred within the social research community as the Solomon Four-Group Quasi-Experimental Design [Frankfort-Nachmias 2008, p. 104]. This design incorporates four experimental treatments, two of which involve a pretest baseline measurement and two that don't. Table 3 depicts this with an example of using four different partner experimental groups (organizational units) that will participate in the experiment.

Table 3: Solomon 4 Group Research Design

Experimental Groups	Pretest (e.g., pre-measurement)	Intervention	Post-test (e.g., post- measurement)
1	Observe1	Intervention	Observe2
2	Observe3		Observe4
3		Intervention	Observe5
4			Observe6

This research design predominantly handles the possibility that baseline measures, taken before experimental conditions (called the *intervention*) are applied, may influence the conduct of the experimental subjects and thus affect the baseline measures taken after the intervention. By incorporating the two treatments without pretest baseline measures, this research design provides greater generalizability by removing any influence of those measurements. If, however, this research determines that there is no danger of influence from the pretest baseline measurements, the design would be reduced to just the first two treatments listed in Table 3 and a straightforward comparison of intervention and lack of intervention. For this research, the intervention will be the purposeful and complete implementation of one or more related resilience practices, while a lack of intervention will be the absence of that implementation. We will implement the Solomon Four-Group Quasi-Experimental Design in an iterative fashion such that each iteration will enable the testing of a different set of resilience practices. Additionally, we will run each iteration for a predefined period of time to observe multiple incidents and the effect of resilience practices on the time and cost required to handle these incidents. Statistical power and sampling rules will be implemented to determine the minimum number of incidents required to conduct a comparison and, consequently, the minimum expected time period required for a given iteration.

Once an iteration of a set of resilience practices is finished, we will conduct a comparative analysis as depicted in Figure 3. Specifically, incident handling times and costs will be recorded for each incident between identified time line events such as

- an incident being reported
- followed by the spread being contained
- followed by system cleansing being completed
- followed by the system(s) being returned to normal operation

Armed with these measurements, we will conduct hypothesis tests to compare the four treatment groups to discern differences in incident handling times and costs for each of the previously identified event segments. Figure 3 depicts the comparison of the different groups graphically with the non-intervention group shown as the red baseline and the intervention group shown as the green improved baseline. Figure 3 also depicts example conclusions from hypothesis tests in which specific differences in incident handling time and cost may be confirmed with different levels of confidence.

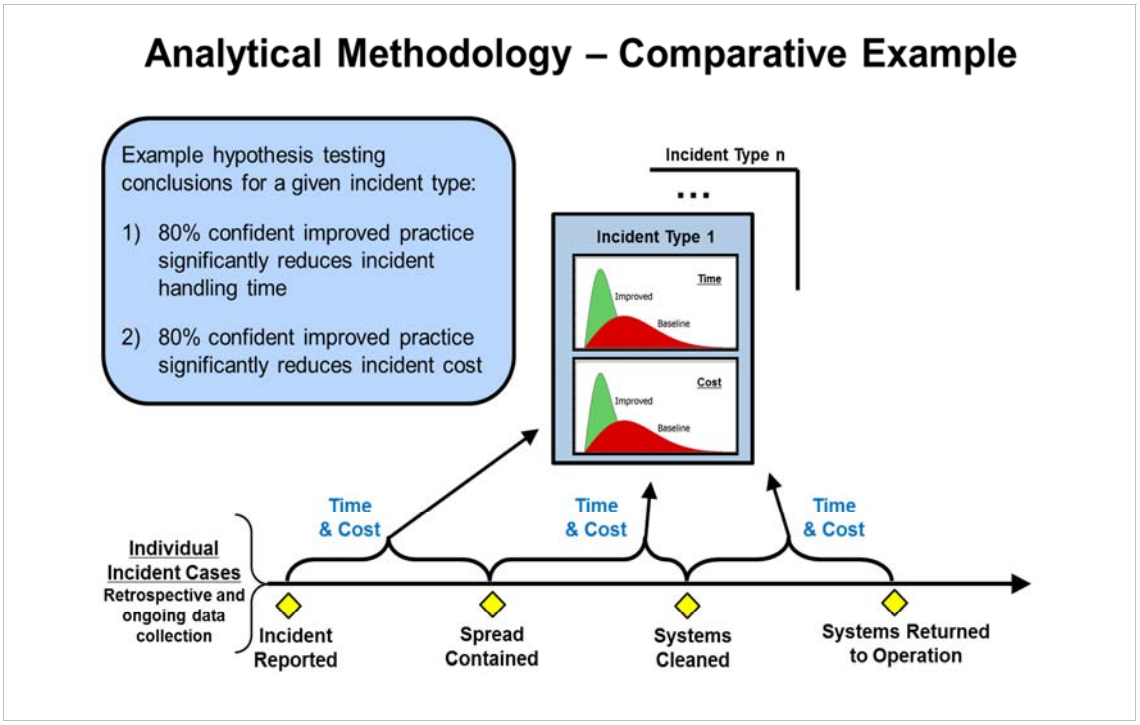

Figure 3: Comparative Analytics of Incident Handling Time and Cost

Once a series of hypotheses are tested for a given set of related resilience practices, as shown in Figure 3, the same data will be used to conduct time-based predictions of incident handling times and costs as depicted in Figure 4. Essentially, Figure 4 demonstrates that the handling times and costs for any given event segment or consecutive segments may be modeled with distribution-fitting approaches, including a modern approach of Weibull modeling [Dodson 1994]. This modeling focuses on the probability that handling times and costs will be below or above (currently arbitrarily) selected thresholds. Figure 4 includes examples of statements that may be concluded from Weibull modeling.

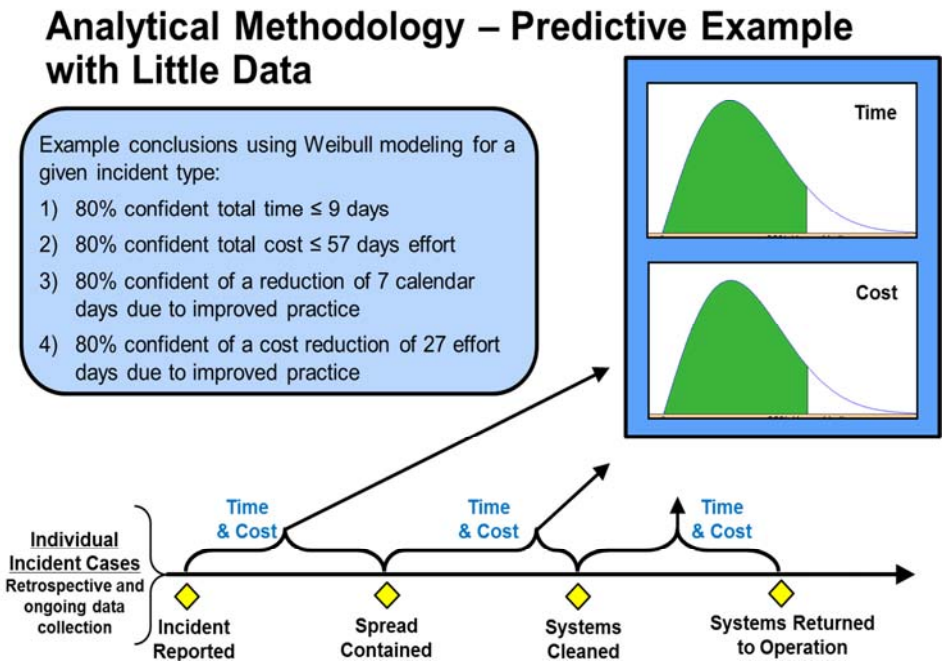

Figure 4: Time-Based Predictive Analytics of Incident Handling Time and Cost

Statements concluded from Weibull modeling may become the basis for establishing a set of benchmarks of incident handling times and costs that can be compared to benchmarks from other organizations or periods of time in the same organization. These time-based predictions may also prove useful in the real-time management of incidents and allocation of resources.

The analytics discussed so far form the rudimentary base of analytics for the resilience practice research. With the collection of more data for each treatment group of a given set of related resilience practices, more sophisticated predictive analytics can be performed as shown in Figure 5. In this predictive analytic approach, measurements from incidents (handling time and cost, implementation status of resilience practices, and other contextual factors) will be used to support statistical regression analysis. The analysis will produce regression equations that predict incident handling time and cost based on the use of specific resilience practices, along with other contextual factors. This approach will enable both a statistical determination of the significance of resilience practices and the actual degree of influence of different resilience practices on handling time and cost.

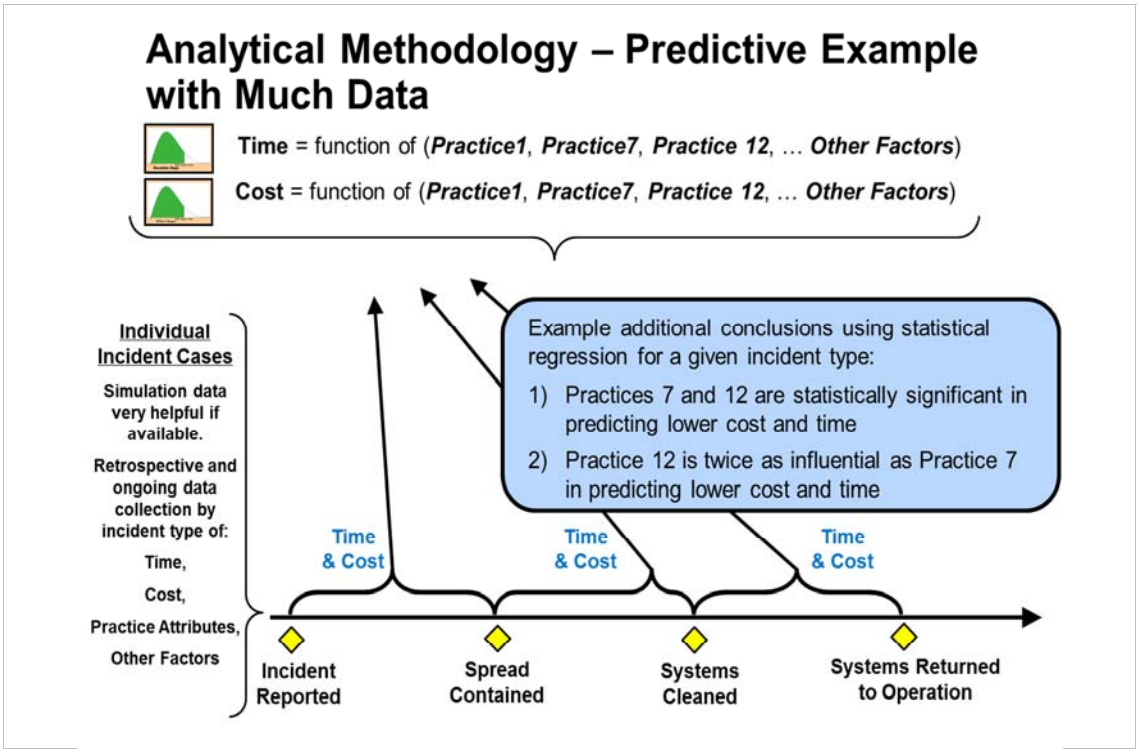

Figure 5: Predictive Analytics of Incident Handling Time and Cost Using Leading Indicators

Regression equations may be developed using data from all organizations and incident types, or just specific segments of them. In either case, the equations will provide predictions with 95% prediction intervals that may serve to support real-time management decision making during an incident. These equations may also be used in a what-if or sensitivity analysis during management planning to help organizations better prepare for handling incidents.

Figure 6 depicts another type of analytical model to pursue during phase 2, one that builds upon probability models commonly used in manufacturing environments. This probability model represents the concept of what is called a first-pass, rolled yield model.

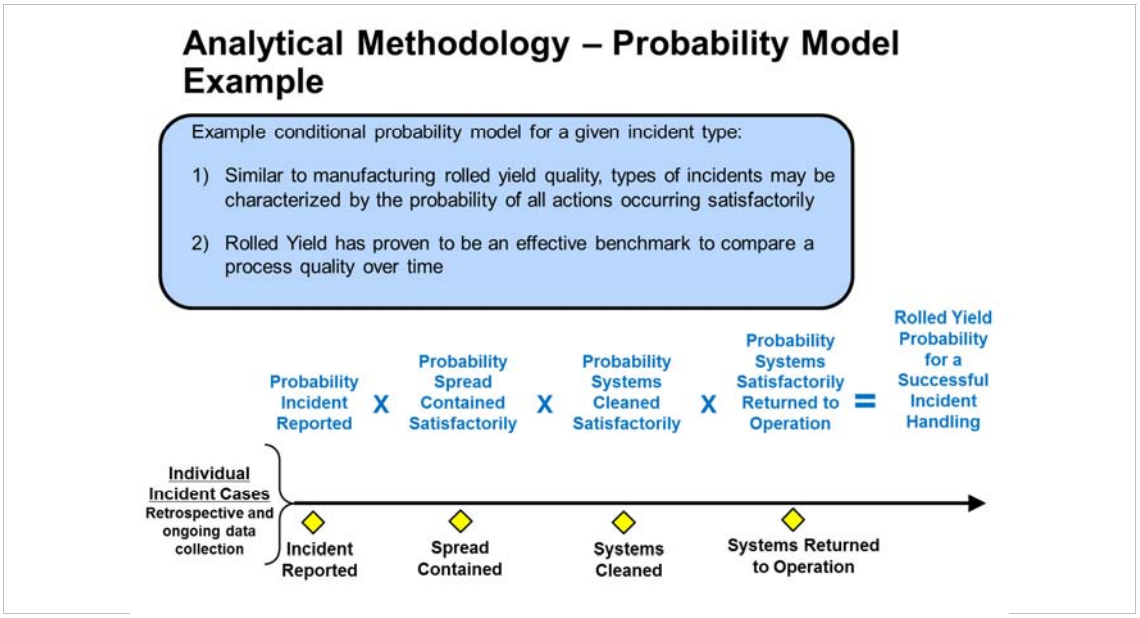

Figure 6: "Rolled Yield" Probability Model of Properly Handled Incidents

As shown in Figure 6, this modeling approach requires a definition of first-pass success for each step in the incident handling, event time line. With such definitions of success, the collected data from the research will enable the computation of conditional success probabilities for each step in the time line. By multiplying these probabilities, a first-pass, rolled yield may be calculated that provides a useful measure for benchmarking and measuring the degree of resilience practice improvement over time for a given organization. Noticeable improvements in resilience practice adoption will produce improved rolled yields for handling incidents.

As this plan indicates, we will use a variety of analytical techniques, beginning with simple analytics that require little measurement data to predictive analytical techniques that are enabled by progressively more research data. This analytical approach provides the research team with more robustness to the various degrees of data to be made available by participating organizations.

4.5 Data Analytics and Validity

The analytical toolkit and research design documented in this plan remain driven by four types of validity:

- operational scope: Are we working on the right problem?
- field significance: Will our solution have the intended impact in the field?
- technical soundness: Is our technical approach scientifically founded?
- technical significance: Is our technical approach well situated in the discipline?

With regard to operational scope, we will use analytics involving designed surveys, metadata analysis, and structured quantitative decision techniques (e.g., the analytic hierarchy process [AHP], multi-attribute utility theory [MAUT], variants of wideband Delphi, and conjoint analysis) to quantify the utility or value of this work to various stakeholders.

With regard to field significance, analytics will play a key role in maximizing external validity through a number of mechanisms. First, the analytical methodology will provide valuable insight into what operational measures should be collected to meet the needs of the resilience performance outcomes and leading indicators. A primary philosophy guiding this activity will come from the research of Douglas Hubbard [Hubbard 2010], who shares a variety of ideas of how to measure intangibles. Measuring intangibles may become a significant activity within phase 2. Second, field quasi-experiments will be designed and analyzed, thereby providing maximum generalizability. Third, analytical results of surveys and structured interviews will provide quantified feedback from stakeholders on the perceived alignment and impact of this research.

With regard to technical soundness, analytics will play a key role in maximizing internal validity by ensuring a sound scientific approach to the research. This analytical plan outlines the use of popular and accepted social science research designs and scientifically driven hypothesis tests whose sampling plans build on sampling units, sampling frames, and sample size determined by acceptable alpha and beta error. Further enhancement of this analytical plan will include the analysis and mitigation of a list of commonly regarded external and internal threats to validity. Each step of the scientific research and analytics will be subject to technical peer review to ensure timely revision of the scientific methods.

With regard to technical significance, analytics will play a key role in contributing new knowledge and innovation within the field of security and resilience. Analytics will be used to quantify the completeness and quality of the early literature search within the resilience domain. Additionally, the research team may choose to take advantage of text analytics as one approach to deciphering large amounts of existing literature and artifacts relevant to this research. Lastly, analytics may also help collaboration partners demonstrate the contribution of this research to their existing research and operational practices.

In summary, this analytical plan describes a set of analytical tools, methods, and activities driven by the four types of validity to ensure a successful first pass at this research in FY13. The analytical rigor described above will also well position the results of phase 2 for distribution in leading research publications and conference presentations.

4.6 Related Work and Communities of Practice

During phase 2, we will identify and compare other approaches being used to answer similar hypotheses and research questions. The primary basis for comparison is the measurement of the impact of applying resilience practices to validate their effectiveness (rather than endorsing the practices based on subject matter expert consensus alone). We will use other criteria as well, such as a focus on compliance to a standard (checklist approach) versus continuous improvement, and the scope of the practices across resilience disciplines—security, business continuity and disaster recovery, and certain aspects of IT operations, all of which CERT-RMM addresses.

Table 4 describes some of the organizations and projects that may be included in our comparison of related work.

Table 4: Prospects for Related Work Comparisons

Entity and Project	Description
Resiliency research, MITRE	A framework of cyber resiliency engineering practices and mapped cyber resiliency metrics. Validation of framework practices through experimentation [Bodeau 2012, Bodeau 2011]
MIT Lincoln Laboratory metrics for DHS Continuous Diagnosis & Mitigation	Security metrics for prevalent network threats; measures of the effect of applied controls (based on the SANS Top 20 Critical Controls) [Lippmann 2011]
SANS Top 20 Critical Controls	20 controls agreed upon by a consortium to be the most effective for preventing and recovering from cyber attacks. Each control is mapped to metrics and measurement guidance [SANS 2011].
NIST 800-53 and NIST 800-55	Guidelines for selecting and specifying security controls for organizations and information systems supporting the executive agencies of the federal government, and guidelines for developing, selecting, and implementing measures to indicate the effectiveness of those controls [NIST 2012, NIST 2008]
ISO 27002 and ISO 27004	Guidelines for initiating, implementing, maintaining, and improving information security management in an organization, and guidance on developing and using measures and measurement to assess the effectiveness of an implemented information security management system and controls or groups of controls [ISO/IEC 2005, ISO/IEC 2009]
CIS Security Metrics	Outcome and practice metrics measuring the frequency and severity of security incidents, incident recovery performance, and the use of security practices that are generally regarded as effective [CIS 2010]
ISO 22301	Requirements for a management system to protect against, reduce the likelihood of occurrence of, prepare for, respond to, and recover from disruptive incidents [ISO 2012]
ResiliNets Architecture, University of Kansas and Lancaster University	An architectural framework for network resilience based on the two-phase strategy D2R2+DR: defend, detect, remediate, recover, diagnose, refine. Ongoing work in simulation, experimentation, and metrics [Sterbenz 2012]
Business Continuity Maturity Model, Virtual Corp.	Open access tool that addresses competencies in business recovery, security management, incident management, technology recovery, and other business continuity program disciplines [Virtual 2012]
Resilience Based Management, U.S. Army Corps of Engineers and others	Continuous resilience management of engineered systems using a recursive process that includes sensing, anticipation, learning, and adaptation [Park 2012]

5 Summary

This report describes a research study that was performed from January through September 2012 (phase 1), and a proposed phase 2 research project to be conducted during FY13 (October 2012 through September 2013), subject to available funding.

The goal of this overall research effort (both phase 1 and phase 2) is to inform the general question "*Do resilience practices have a discernible (measurable) effect on operational resilience?*" In the FY13 phase 2 proposal, we express the objective as a statement: *Measure the contribution of resilience practices to reducing the occurrence and impact of disruptive events.*

Our general hypothesis is

> *Disruptive events are more likely to have a positive outcome (i.e., response considered successful) when resilience practices are implemented than when they are not.*

After presenting our objectives, research questions, and hypotheses, we describe the various directions we pursued to define resilience cases, engage with collaboration partners, collect case data, and attempt to analyze such data. We share a wide range of observations and insights gained during phase 1 and present a summary of the proposed phase 2 project. That summary includes an in-depth discussion of promising analysis approaches and a first look at other related research efforts that we intend to build upon.

The authors of this report welcome readers' comments and questions. We are actively seeking collaboration partners who may be willing to share their incident data with us and receive the results of our ongoing research in return.

Appendix Resilience Case Request for Data Template

Analyze Cases of Resilience Success and Failure

Request for Data: Please provide a description of organization events[15] where resilience processes/practices provided, or could have provided, more effective resilience in the form of protecting and sustaining assets (e.g., people, information, technology, and facilities) so that the organization can continue to carry out its mission.

Organization Information	Description
Sector	Is the organization within private or public sector? If public sector, indicate whether federal, state, or local government. If private sector, indicate the vertical domain (e.g., financial, aerospace, defense, energy, academia, non-profit, etc.)
Size	Size of the organization (measured by one or more of quantities such as annual revenue or the number of employees, contractors, subcontractors, suppliers, or others granted authorized access to critical assets)
U.S. Organization	U.S. Only Organization; (Y)es, (N)o: Does the organization conduct operations solely in the U.S.?
U.S. State	The state in the U.S. where the organizational unit, impacted by the event, is located.
Country	If the organizational unit, impacted by the event, is outside the U.S., indicate the country where the organizational unit is located.
Deployed resilience/security frameworks and models	ISO 27000x, ITIL, NIST, CMMI, CERT-RMM, etc. (those relevant to the scope of CERT-RMM)
Length of time deployed framework in place	How long prior to the event was the deployed resilience/security framework in place?
Sources	List all data sources used to fill out this template. If there are multiple sources, mark them as [1], [2], [3], etc. with normal citation information and then use the bracketed numbers in each data field to identify the source(s) of each data item. If most or all of the information is from the same source, list it as the first source. In that case, no need to use the bracketed [1] in the data fields; any unmarked field will be assumed to be based on information from the first citation.
Date initial case data provided	Date when the resilience case analysis research team received the first description of the case

[15] Event - one or more occurrences that affect organizational services and assets, and have the potential to disrupt operations

Event Information	Description
Summary	Anonymized and fairly high-level description of the event
Begin Date	Date the event was first detected (or the first impact of the event)
Chronology	Brief description of event time line and key actions taken
Actions Taken	Actions taken in response to the event
Date of Impact	The date the organization was impacted by the event, which rendered a critical system or service unavailable or degraded. The date of impact may be different than the date the organization observed that a critical system or service was disrupted.
Date Observed	The date the organization observed or detected that a critical system or service was disrupted. The date observed may be different than the date the organization was impacted by the event.
Date Addressed	The date the organization *began* taking action to restore the critical system or service to normal operating capacity.
Date Restored	The date the organization *finished* taking action to restore the critical system or service to normal operating capacity.
Date Closed	Date the event was considered closed
Duration	The length of time, measured in hours, days, weeks, months, or years, that the critical system or service was not available due to the impact of the event
Source of Event	Internal or external, i.e., was the event caused by an internally initiated action (e.g., rogue employee) or an externally initiated action (e.g., an external attacker)
Financial Consequence - Downtime	Financial impact associated with the event due to disrupted service(s), system(s), or other asset(s), i.e., the cost associated with allowing the organization to continue operations, possibly in a reduced state. Examples: the cost of executing a business continuity plan or the cost of using a secondary server
Financial Consequence – Incident Resolution	Financial impact associated with the event due to cost expenditures on incident resolution
Financial Consequence – Revenue Loss	Revenue lost as a result of the event, i.e., the revenue not generated by the organization as a result of the disrupted service, system, or asset
Financial Consequence - Other	Other financial impacts, not listed above, associated with the event
Operational Consequence – Disrupted Services	Services disrupted and impact of service disruption as a result of the event
Operational Consequence – Personnel Impacts	Staff effort required to handle/recover from the event, by role if possible
Operational Consequence – Additional Resources	Resources consumed in handling the event in addition to staff time (facilities, technology, etc.)
Operational Consequence – Other	Other operational impacts, not listed above, associated with the event
Vulnerabilities	Exposures, flaws, or other weaknesses that made the affected organization susceptible to the event
Key Processes	Processes that were essential to successful handling of the event (impact within established tolerances/threshold)
Key Practices	Practices that were essential to successful handling of the event (impact within established tolerances/threshold)
Practice[16] Gaps	Practices that would have reduced impact had they been implemented or more effective
Process[17] Gaps	Processes that would have reduced impact had they been implemented or more effective

[16] Practice: a method or technique for accomplishing some objective

[17] Process: activities that can be recognized as implementations of practices

Event Information	Description
Available for Interview	Are you available for an interview by the research team to allow for additional fact-finding questions regarding the event? (Y)es; (N)o
Organization Collaborator	Organization supplying the case information
Individual Collaborator	Individual supplying the case information
Case Coder	Research team member(s) who coded this case

Glossary

consequence
The unwanted effect or undesirable outcome on the organization as the result of exploitation of a condition or threat (CERT-RMM RISK). There is a consequence if an Impact > Threshold for some resilience requirement. An example of a consequence is the lack of availability of a key customer facing website for 48 hours due to malware infection. As a result, customers are unable to complete specific business transactions leading to a measurable reduction in revenue for the outage time period—an expression of impact.

effort to restore
The total number of staff hours that are required to restore operational capability during and following a disruptive event (incident).

event
One or more occurrences that affect organizational assets and have the potential to disrupt operations (CERT-RMM IMC).

impact
To have a direct effect upon (www.merriam-webster.com); a level of productive capability of an asset that has been lost as the result of exploitation of a condition or threat, after all incident response actions have been taken; expressed as time to restore, effort to restore, and cost to restore for phase 1.

incident
An event (or series of events) of higher magnitude that significantly affects organizational assets and requires the organization to respond in some way to prevent or limit organizational impact (CERT-RMM IMC).

incident response
The actions the organization takes to prevent or contain the impact of an incident to the organization while it is occurring or shortly after it has occurred (CERT-RMM IMC).

resilience case
A case in which an organization was successful or unsuccessful in satisfying a resilience requirement in the presence of and after a disruptive event. A resilience case may involve either a consequence or resilience evidence.

resilience evidence
Demonstrated resilience of an organization as the result of exploitation of a condition or threat. Resilience evidence accrues if an Impact <= Threshold for some resilience requirement. An example of resilience evidence is service and personnel transition to a backup facility with restored, core IT operations and telephony services within 24 hours (threshold) as specified in the service continuity plan.

resilience practice
A method or technique that helps satisfy a resilience requirement.

resilience requirement
A constraint that the organization places on the productive capability of an asset (information, technology, facilities, personnel) to ensure that it remains viable and sustainable when charged into production to support a service (CERT-RMM). Availability service level agreements for key servers are examples of a resilience requirement.

resilience threshold (or just threshold)
The minimal level of productive capability of an asset as determined by resilience requirements.

sustainment practices
Activities and use of related controls necessary to maintain an asset in a desired operational state when it is subjected to harm or disruptive events.

time to restore
The total number of hours, days, weeks, or months that are required to restore operational capability during and following a disruptive event (incident); this may also include total elapsed (calendar) time.

References

URLs are valid as of the publication date of this document.

[Allen 2011a]
Allen, Julia & Curtis, Pamela. *Measures for Managing Operational Resilience* (CMU/SEI-2011-TR-019). Software Engineering Institute, Carnegie Mellon University, 2011. http://www.sei.cmu.edu/library/abstracts/reports/11tr019.cfm

[Allen 2011b]
Allen, Julia; Curtis, Pamela; & Gates, Linda. *Using Defined Processes as a Context for Resilience Measures* (CMU/SEI-2011-TN-029). Software Engineering Institute, Carnegie Mellon University, 2011. http://www.sei.cmu.edu/library/abstracts/reports/11tn029.cfm

[Allen 2010]
Allen, Julia & Davis, Noopur. *Measuring Operational Resilience Using the CERT Resilience Management Model* (CMU/SEI-2010-TN-030). Software Engineering Institute, Carnegie Mellon University, 2010. http://www.sei.cmu.edu/library/abstracts/reports/10tn030.cfm

[Bodeau 2012]
Bodeau, Deborah J.; Graubart, Richard; LaPadula, Len; Kertzner, Peter; Rosenthal, Arnie; & Brennan, Jay. *Cyber Resiliency Metrics,* Version 1.0, Rev. 1. The MITRE Corp., April 2012. https://register.mitre.org/sr/12_2226.pdf

[Bodeau 2011]
Bodeau, Deborah J. & Graubart, Richard. *Cyber Resiliency Engineering Framework* (MTR110237). The MITRE Corp., September 2011. http://www.mitre.org/work/tech_papers/2012/11_4436/

[Caralli 2011]
Caralli, Richard A.; Allen, Julia H.; & White, David W. *CERT Resilience Management Model: A Maturity Model for Managing Operational Resilience*. Addison-Wesley, 2011. http://www.sei.cmu.edu/library/abstracts/books/9780321712431.cfm

[Caralli 2010]
Caralli, Richard A.; Allen, Julia H.; Curtis, Pamela D.; White, David W.; & Young, Lisa R. *CERT® Resilience Management Model, V1.0 - Process Areas, Generic Goals and Practices, and Glossary (CERT-RMM v1.0).* Software Engineering Institute, Carnegie Mellon University, 2010. http://www.cert.org/resilience/rmm.html

[CERT 2012]
CERT Program, Software Engineering Institute. *CERT Resilience Management Model.* http://www.cert.org/resilience/rmm.html (2012).

[CIS 2010]
Center for Internet Security (CIS). *The CIS Consensus Information Security Metrics*. http://benchmarks.cisecurity.org/en-us/?route=downloads.metrics (November 2010).

[Creswell 2011]
Creswell, J. W. & Clark, V. L. P. *Designing and Conducting Mixed Methods Research*, 2nd Edition. Sage Publications, 2011.

[Dodson 1994]
Dodson, B. *Weibull Analysis*. American Society for Quality, 1994.

[Frankfort-Nachmias 2008]
Frankfort-Nachmias, C. & Nachmias, D. *Research Methods in the Social Sciences*, 7th ed. Worth, 2008.

[Hubbard 2010]
Hubbard, D. W. *How to Measure Anything*, 2nd ed. Wiley. 2010.

[ISO 2012]
International Organization for Standardization (ISO). *Societal security—Business continuity management systems—Requirements* (ISO/IEC 22301:2012). ISO, 2012.

[ISO/IEC 2009]
International Organization for Standardization/International Electrotechnical Commission (ISO/IEC). *Information technology—Security techniques—Information security management—Measurement* (ISO/IEC 27004:2009). ISO/IEC, 2009.

[ISO/IEC 2005]
International Organization for Standardization/International Electrotechnical Commission (ISO/IEC). *Information technology—Security techniques—Code of practice for information security management* (ISO/IEC 27002:2005). ISO/IEC, 2005.

[King 2011]
King, Steven E. *Cyber S&T Priority Steering Council Research Roadmap*. National Defense Industrial Association Disruptive Technologies Conference, November 2011. http://www.acq.osd.mil/chieftechnologist/publications/docs/2011%2011%2001%20Cyber%20PSC%20Roadmap.pdf

[Lippmann 2011]
Lippmann, R. P.; Riordan, J. F.; Yu, T. H.; & Watson, K. K. *Continuous Security Metrics for Prevalent Network Threats: Introduction and First Four Metrics* (Project Report IA-3). Lincoln Laboratory, 2012.

[Montgomery 2012]
Montgomery, D. C. *Design and Analysis of Experiments*, 8th ed. Wiley, 2012.

[NIST 2012]
National Institute of Standards and Technology. *Security and Privacy Controls for Federal Information Systems and Organizations* (NIST 800-53), Revision 4, Initial Public Draft. NIST, 2012. http://csrc.nist.gov/publications/PubsSPs.html

[NIST 2008]
Chew, Elizabeth; Swanson, Marianne; Stine, Kevin; Bartol, Nadya; Brown, Anthony; & Robinson, Will. *Performance Measurement Guide for Information Security* (NIST 800-55), Revision 1. NIST, 2008. http://csrc.nist.gov/publications/PubsSPs.html

[Park 2012]
Park, J.; Seager, T. P.; Rao, P. S. C.; Convertino, M.; & Linkov, I. "Integrating Risk and Resilience Approaches to Catastrophe Management in Engineering Systems." *Risk Analysis* DOI: 10.1111/j.1539-6924.2012.01885.x (September 11, 2012). http://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2012.01885.x/abstract

[Partridge 2011]
Partridge, Kevin & Young, Lisa. *CERT Resilience Management Model (CERT-RMM) V1.1: Code of Practice Crosswalk Commercial Version 1.1* (CMU/SEI-2011-TN-012). Software Engineering Institute, Carnegie Mellon University, 2011. http://www.sei.cmu.edu/library/abstracts/reports/11tn012.cfm

[SANS 2011]
SANS Institute. *Twenty Critical Security Controls for Effective Cyber Defense: Consensus Audit Guidelines, V3.1*. SANS Institute, October 3, 2011. http://www.sans.org/critical-security-controls/

[Sheskin 2011]
Sheskin, David J. *Handbook of Parametric and Nonparametric Statistical Procedures*, 5th ed. Chapman & Hall/CRC, 2011.

[Sterbenz 2012]
Sterbenz, James P. G. et al. *ResiliNets: Resilient and Survivable Networks*. University of Kansas and Lancaster University, 2012. https://wiki.ittc.ku.edu/resilinets/Main_Page

[Virtual 2012]
Virtual Corp., Inc. *The Business Continuity Maturity Model*. http://www.virtual-corp.net/html/bcmm.html (2003).

[Yin 2009]
Yin, R. K. *Case Study Research: Design and Methods*, 4th ed. Sage Publications, 2009.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
(Leave Blank)	December 2012	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Analyzing Cases of Resilience Success and Failure—A Research Study	FA8721-05-C-0003

6. AUTHOR(S)

Julia H. Allen, Pamela Curtis , Nader Mehravari, Andrew Moore, Kevin Partridge, Robert Stoddard, & Randy Trzeciak

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	CMU/SEI-2012-TN-025

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
SEI Administrative Agent AFLCMC/PZE 20 Schilling Circle, Bldg 1305, 3rd floor Hanscom AFB, MA 01731-2125	n/a

11. SUPPLEMENTARY NOTES

12A DISTRIBUTION/AVAILABILITY STATEMENT	12B DISTRIBUTION CODE
Unclassified/Unlimited, DTIC, NTIS	

13. ABSTRACT (MAXIMUM 200 WORDS)

Organizations that are using the CERT® Resilience Management Model and organizations that are considering using it want information about the business value of implementing resilience processes and practices, and how to determine which ones to implement. This report describes the SEI research study that begins to address this need. It includes a discussion of the completed phase 1 study and a proposed phase 2 project. Phase 1 included forming a hypothesis and set of research questions and using a variety of techniques to collect data and evaluate whether resilience practices have a discernible (measurable) effect on operational resilience—that is, an organization's ability to continue to carry out its mission (provide critical services) in the presence of operational stress and disruption. The outcomes of phase 1 provide the foundation for the proposed phase 2. The longer term goal includes developing a quantitative, validated business case for prioritizing and implementing specific resilience practices, including decision criteria for selecting and measuring investments in improved resilience.

14. SUBJECT TERMS	15. NUMBER OF PAGES
security, model, RMM, resilience	50

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL